



Secure Access to Administrative Data in France Context, *Governance and Technology*

Roxane Silberman (CASD-GENES, and CNRS Paris)

*Commission on Evidence-Based Policymaking meeting, US
January 13, 2017*

- **Key points in the legal frameworks context**
- **The overall process for access**
- **Governance**
- **CASD role and technological solution**

INTRODUCTION

- Opening access to confidential government microdata for research in France : a 2 steps process
 - Access to official confidential microdata (census and surveys) 2008
 - Access to administrative data in France since 2014-2016
- Administrative data under a rather dispatched organisation and legal frameworks
- However organization of access to administrative data has benefited from the centralized process and infrastructure set up for access to official microdata
 - The approval process: *Comité du secret statistique*
 - The remote access infrastructure: *CASD*

KEY POINTS IN THE LEGAL FRAMEWORKS FOR ADMINISTRATIVE DATA

- Under various legal frameworks
- Started by changing specific legal frameworks one by one
 - 2014 Tax data
 - 2015 Medico-administrative data
- Extension to all administrative data - 2016
 - By clearing of the responsibility of producers regarding professional secrecy
 - NB Central Bank data depend on the European framework: Banque de France opened an on-site access
- Last step: Facilitate linkage on the basis of the NIR (National Identity Number) for statistical and research purpose - 2016

THE OVERALL PROCESS

- A researcher can ask any data under the legal perimeter
- The researcher contacts the producer to check data fit the research project
- Approval process via independent authorities
+ Agreement of the producer (see above)
- Remote access process: CASD acts as third party between the producers and the users

GOVERNANCE (1): THE APPROVAL PROCESS

- *Comité du secret statistique*
 - An independent authority chaired by a lawyer
 - Representatives of the Parliament
 - Representatives of the Unions
 - Representatives of the producers
 - Representatives of CNIL (National Authority in charge of privacy protection) and of Archives de France (National Archives)
 - 2 Representatives of researchers (proposed by the ministry of research in consultation with the Data Archives infrastructure)
- Provides an advice based on the research project and public interest
 - + Agreement of the producer
 - + Formal approval from the Authority in charge of privacy protections if personal data (CNIL)
 - + Formal approval from the National Archives
- Does not require accreditation of institutions
- Open to researchers from European Union

GOVERNANCE (2): THE SECURE ACCESS CASD

- A department of GENES, an independent institution of high education
 - Representatives of Ministry of research and Ministry of finances on board
- CASD governing board includes 5 Founding Research institutions + Data archive infrastructure + INSEE
- A Scientific Advisory Board
 - 15 members from France and other countries
 - Researchers from various disciplines, IT and methodologists + a representative of INSEE, Eurostat invited

THE ACCESS PROCESS: CASD ROLE

- Acts as third party for secure access between data producers and researchers
- Equal access for all researchers (no specific access for the founding institutions)
 - Preparation and transfer of data with the producers
 - Metadata harmonised under DDI
 - Establish contracts with researchers (responsibility of the researcher)
 - Enrolment and training of researchers
 - Secure remote access to data
 - Outputs checked before export or automatic exports depending on agreements with data holders
 - Can act as third party for data linkage

CASD TECHNOLOGICAL SOLUTION

A distributed architecture designed by GENES IT department

- **High security**

- An integrated system
- Physical and logical security including a biometric strong authentication
- No possibility to download data

- **Usability for the researchers**

- Remote access: researchers can work from their university with their SD-Box (+ access card +biometric control)
- Researchers can see the data
- A research environment for working and writing

- **Flexibility**

- Easy to install and deploy in the pre-existing researchers' institutions environment: no compatibility issues, no test to run

CASD : an integrated system

:

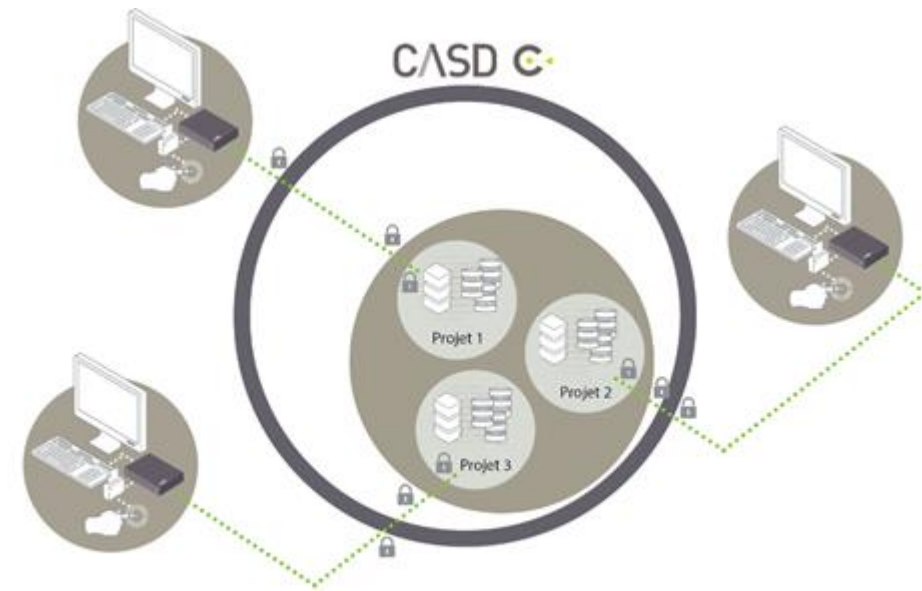
- A central IT infrastructure highly secure
- The SD-Box : the only way to access the central infrastructure. They are sent to users.



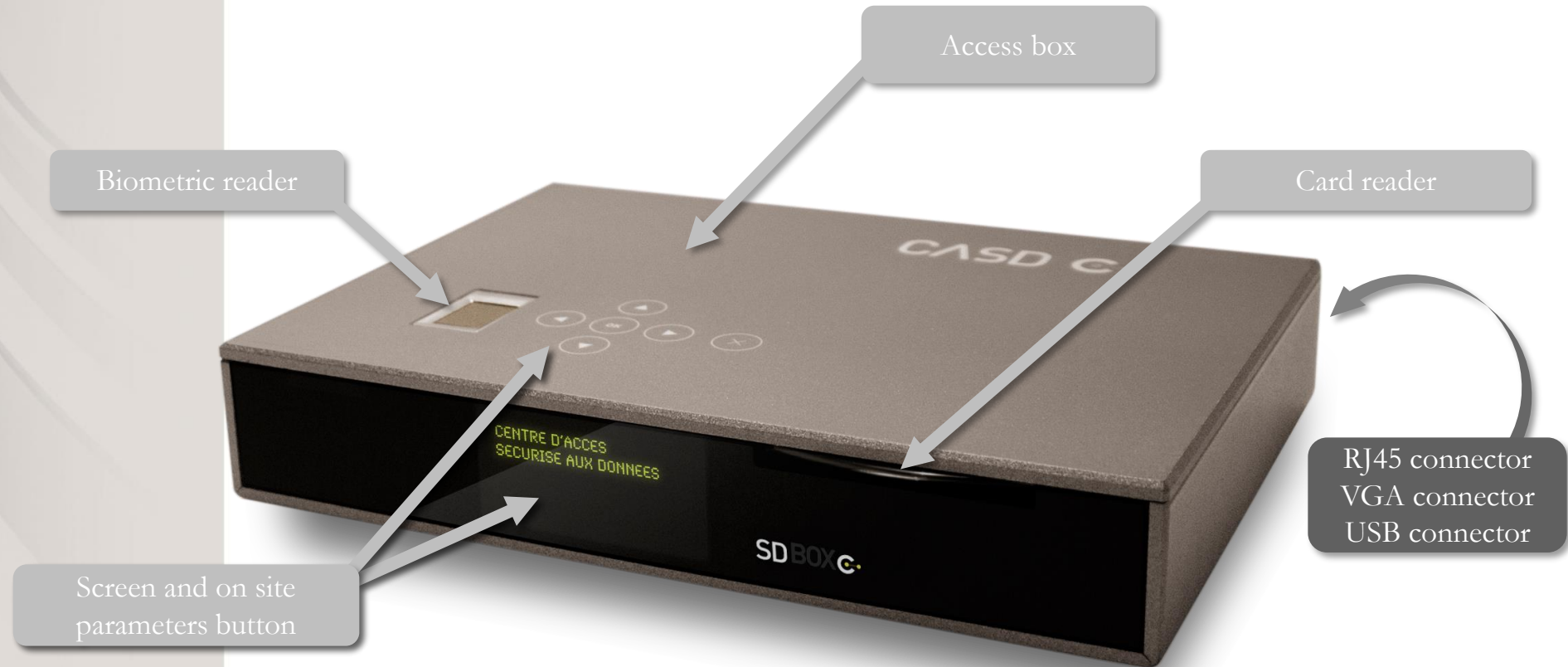
GENES

GRUPE DES ÉCOLES NATIONALES D'ÉCONOMIE ET DE STATISTIQUE

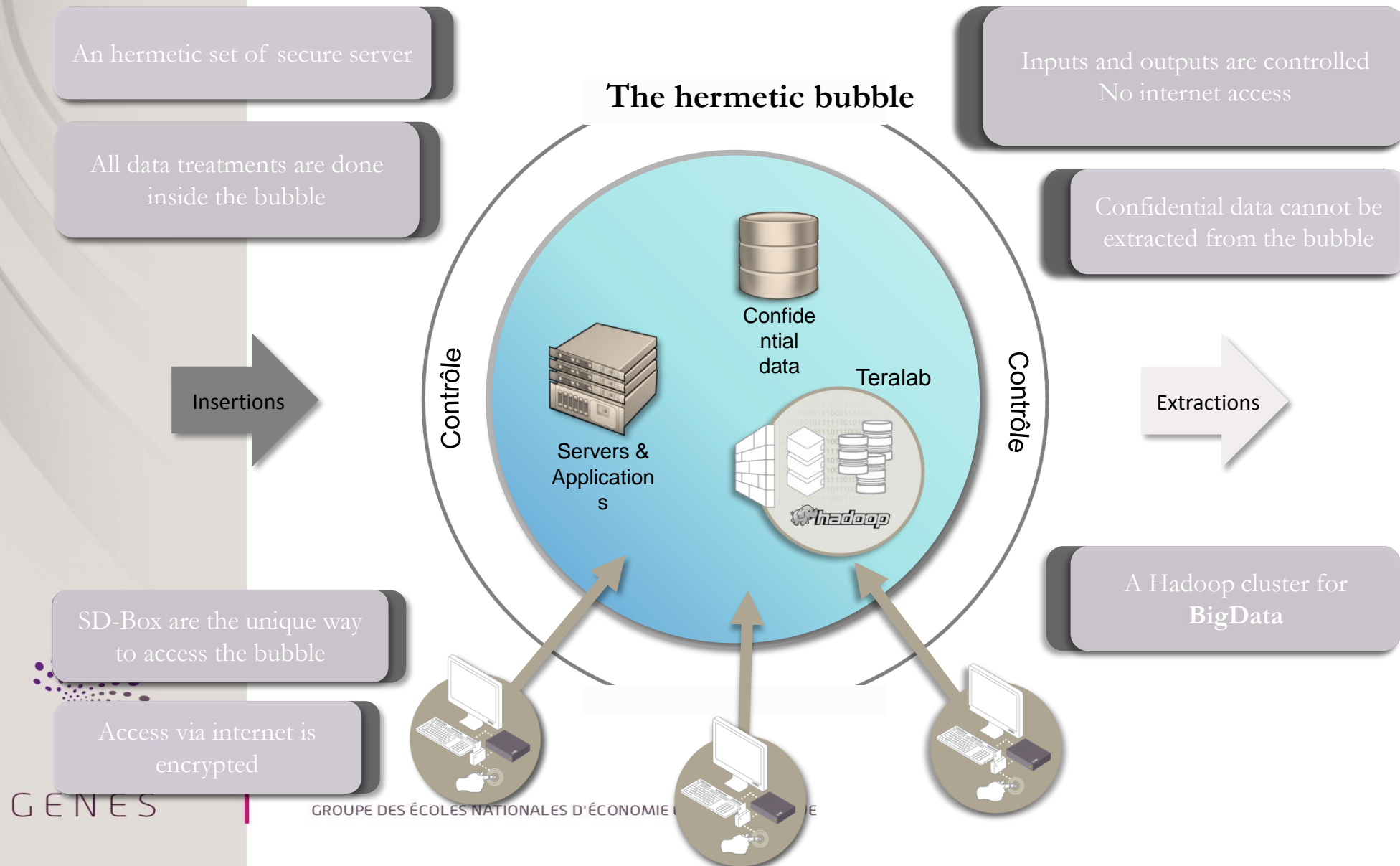
La SD-Box.



The SD-Box®



The bubble: an hermetic infrastructure



CONCLUSION

The technological solution:

- **From the researcher's perspective:** everything is at hand, data, a well-known environment with all software, storage, computing power, also designed for big data
- **From the security perspective:** everything happens at CASD in the bubble with no dependence of other systems that may compromise security, no possibility to download data
- **From a service perspective (installation, supervision, assistance)** SD-Box are identical, monitored, configured, updated remotely by CASD, easy to replace
- **From a cost perspective:** the solution integrates software and services. SD-Box can be shared by several teams

The overall unified system :

- **A central point for access with harmonised criteria for approval, same procedures for access and possibility to easily work with datasets from different administrations**



Sommaire

THANKS !

roxane.silberman@casd.eu

<https://casd.eu>